2013

A Performance Metric and Goal-Setting Procedure for Deadline-Oriented Processes

Doerr, Kenneth Howard

Production and Operations Management

þÿProduction and Operations Management (POMS) Vol. 22, No. 3, May June

PRODUCTION AND OPERATIONS MANAGEMENT

PRODUCTION AND OPERATIONS MANAGEMENT

Vol. 22, No. 3, May–June 2013, pp. 726–738
ISSN 1059-1478 | EISSN 1937-5956 | 13 | 2203 | 0726
DOI 10.1111/j.1937-5956.2012.01375.x

A Performance Metric and Goal-Setting Procedure for Deadline-Oriented Processes

Kenneth Howard Doerr
Graduate School of Business & Public Policy, Naval Postgraduate School, Monterey, California 93943, USA, khdoerr@nps.edu

Kevin R. Gue
Dept of Industrial & Systems Engineering, Auburn University, Auburn, Alabama 36849-5346, USA, Kevin.Gue@auburn.edu

A performance metric and goal-setting procedure is defined for an order fulfillment operation. In this operation, order requests arrive continuously, and filled orders are shipped at a specific time each day. The metric links the continuous operation of order fulfillment to the scheduled shipment times. To prescribe goals against the metric, a performance model is developed that incorporates the motivational effect of a goal. Goal-Setting Theory is used to establish the performance goal and to show how to match arriving orders to deadlines based on their arrival times and expected processing times. Monte Carlo simulation on data from a large distribution center is used to demonstrate that setting these two parameters in the light of motivational research yields quite different results than doing so with an intuitive method. Moreover, a motivational goal leads to better operational performance; that is, correctly setting up the metric causes more customers to receive their orders sooner.

Key words: metrics; goals; warehousing; deadlines; motivation
History: Received: August 2009; Accepted: March 2012 by Karen Donohue, after 3 revisions.

1. Introduction

Order fulfillment is at the heart of a consumer-based service economy. For retailers in particular, the ability to respond quickly to customer orders can mean the difference between delighted and disappointed customers. Conceptually, order fulfillment is the straightforward process of picking products off of shelves, putting them in boxes, and then loading them onto trucks. The reality is much more complex: hundreds of workers must coordinate their activities in warehouses as large as 1 million square feet to deliver thousands of items to the shipping dock before the trucks leave each day. Missing the truck departure by just 5 minutes probably means a delay to the customer of 24 hours.

This research is focused on improving customer service in such operations by developing a metric and goal-setting procedure that links the continuous processes of picking and packing orders to the scheduled, batch process of truck departures. The research is motivated by the order fulfillment operations in a large distribution center in California.

At this distribution center, orders arrive around the clock, but shipments leave the facility only at specified times late in the day. Many shipments leave via UPS or FedEx; others leave via less-than-truckload carriers. Customers of this distributor expect short response times to orders. Service performance at the DC was internally assessed by *average flow time*, measured as the average elapsed time between an order being received and being ready for shipment. The corporate goal was to have this average be shorter than 24 hours. Unfortunately, using this internal measure of flow time induced some unintended behavior. For example, managers often invoked overtime on Friday nights to clear orders for shipment, even though no shipments were scheduled over the weekend. This behavior was consistent with the goal of improving performance by reducing *average flow time*, even though it provided no benefit for the system (in fact, it incurred additional costs).

An alternative metric for order fulfillment operations, which better aligns incentives with system level performance, is *Next Scheduled Deadline* (NSD). NSD is defined as the percentage of orders targeted for a particular deadline (truck departure) that are ready by that deadline. Orders are targeted for particular deadlines through the use of a cutoff time: each deadline having an associated cutoff time. Orders arriving before the cutoff time are targeted for the associated deadline.

NSD has several attractive attributes that companies should look for in an effective performance measure. First, for a given cutoff time, an increase in the measure corresponds to a direct improvement in customer service—some customers receive their packages a day earlier than they otherwise would. Second, the measure incorporates the non-continuous nature of shipping and motivates workers to work at an accelerated pace when it matters most—just before the deadline. Third, the measure is meaningful and easy to communicate to customers. This third attribute allows the measure to be used both internally and externally, albeit with different target levels. (In this paper, only the internal measure is discussed.)

Using the NSD measure as part of a motivational goal requires setting both the NSD percentage target and the associated cutoff time. The goal of our research was to examine how these two levers should be set to improve customer service. This question is examined by developing a performance model from the tenets of Goal-Setting Theory, then using Monte Carlo Simulation and a bootstrapping procedure to evaluate potential solutions and set the two parameters. The procedure is demonstrated with data from the field site. The procedure is heuristic in that it does not perform an exhaustive search of all possible parameter settings (in particular, cutoff times are only examined in one-hour increments), but a *post hoc* analysis is used to examine the quality of the solution.

Results show that using well-supported results from Goal-Setting Theory leads to parameter settings (an *operational* policy) that are different than they otherwise would be. In particular, nominal targets are set lower and cutoff times set later than an intuitive process would suggest. Extrapolating from existing goal-setting research findings, analysis of the field data shows that 20.2% more orders would be placed on the next shipment by setting parameters with the procedure described in this paper, rather than an intuitive procedure.

It is important to note that the procedure described below *uses* existing results from Goal-Setting Theory to build a prescriptive model for operational improvement. That is, the data are not analyzed to support (yet again) the results predicted by Goal-Setting Theory. Rather, the data are used to set an operational policy, assuming that Goal-Setting Theory is correct in its main predictions.

A review of the relevant literature in operations management, performance measurement, and goal setting follows in section 2. Section 3 introduces a parsimonious performance model that incorporates behavioral effects and also introduces a goal-setting procedure. Section 4 applies the model to the field data and demonstrates the superiority of the proposed procedure to an intuitive one. The paper concludes in section 5 with a discussion of managerial implications, limitations, and ideas for future research.

2. Literature Review

At least since the publication of Relevance Lost (Johnson and Kaplan 1987), there has been an acute awareness of the sometimes dysfunctional connection between performance measures and an organization's ability not only to assess but also to meet its tactical and operational objectives. Recent performance measurement research has increasingly focused on both financial and non-financial measures, especially those which help employees understand how their work meets customer needs (Euske and Zander 2005). Metrics not only measure but also change organizational performance. At times, performance measurement systems encourage employee behavior that does not advance operational objectives (Kerr 1975).

One of the primary gaps between the operations perspective and the financial perspective has to do with the impact and use of time (Neely and Austin 2002). There is a need for prospective or at least real-time measures of performance vs. the traditionally backward-looking historical financial measures. There is also a need for the proper accounting of the utilization of time itself, as time is connected with both corporate strategy and customer satisfaction (Stalk and Hout 1990). Managers increasingly incorporate operational metrics to ensure that performance measurement encourages the right behavior, many of which are deadline oriented, for example, "critical team objectives like filling an order within 24 hours" (Meyer 1998).

Motivation, of course, is not in the domain of management accounting, but rather organizational behavior. In reviewing the literature on motivation, Goal-Setting Theory has a central place (Mitchell and Daniels 2001). It is also an important area in which OB research can inform operational models (Bendoly et al. 2010, Boudreau et al. 2003). In Goal-Setting Theory, the effect of a goal on performance is *mediated* by motivation. That is, setting a goal only affects performance *through* the impact it has on motivation. The mechanisms of this mediation have been the subject of considerable research and are fairly well understood (Locke and Latham 1990). Goals, when accepted (i.e., internalized), improve performance through behaviors such as attention, effort, persistence, and improved task strategies. Goals also encourage innovation by encouraging the development of improved strategies for accomplishing a task. Thus, there are many ways that goal setting, through motivation, can increase performance.

The basic finding that specific, difficult goals have a positive impact on performance has been validated numerous times through original work and several meta-analyses (Johnson et al. 1981, Tubbs 1986, Wood et al. 1987). The critics who do exist tend to be long on anecdote and short on data (Locke and Latham 2009). The goal-setting effect is arguably the best-known result from the last 25 years of organizational behavior research into motivation. Nonetheless, the impact of Goal-Setting Theory has not been felt in all areas of work or academia, and its prescriptions are not necessarily embedded in industrial systems. Moreover, to some extent, operations management models may have overlooked goal-setting prescriptions because they are difficult to model and are not as precise and simple as they seem at first glance.

This is not to say that the motivation and goal-setting literature has been totally ignored by Operations Management researchers. Within the domain of Social Psychology, Bendoly et al. (2010) list motivation and goal setting as a primary body of knowledge supporting the emerging field of Behavioral Operations. Hays and Hill (2001, 2006) examined motivation as a mediator between service guarantees and service quality. Doerr et al. (1996) investigated the effect of group goals vs. individual goals in push vs. pull production. Linderman et al. (2003, 2006) observed the role of goals in six-sigma process improvement teams and found that difficult goals can sometimes be counterproductive. Bendoly and Prietula (2008) looked at the well-known interaction of training and goals (Mitchell et al. 1994) and discovered further support for the hypothesis that difficult goals can be counter-productive during skill acquisition. Still, these papers are the exception to the rule: although goals are pervasive in Operations, goal-setting research is rarely drawn upon for support, and findings have rarely been embedded in prescriptive models, as they are in this paper. One exception to this is Yuen (2006), who developed a prescriptive model of quality/quantity trade-offs in white-collar workers given deadline-based goals.

Again, the relative dearth of Operations Management papers drawing on goal-setting research might be explained by a great number of complications to the seemingly simple idea that specific, difficult goals should improve performance. Some of these complications are fundamental: for example, how “difficult” should the goal be, and by how much will performance improve? But the theory has also been extended, and many moderators (factors that may modify the basic relationship among goals, motivation, and performance) have been implicated. Some of these have been researched quite deeply, like whether participative goals work better than assigned goals, but others have received relatively less attention. For example, how does performance change over time in the presence of a goal that must be attained by a deadline?

The fact that time is central to the motivational effect of goals is not in question. Another motivated behavior shown to arise from goals is persistence over time (Locke and Latham 1990). Recent work has elaborated on this by modeling persistence as a factor that reduces instability between intentions and actions over time (Meier and Albrecht 2003). One early meta-analysis (Tubbs 1986) found that goals specifying quantity had a stronger effect ($d = 0.845$) than those specifying a time by which the task was to be completed ($d = 0.420$), indicating that the goal-setting effect is sensitive to the clock. Other work (Mitchell et al. 2008, Vancouver et al. 2001, 2005) has examined the dynamic nature of motivation and goal-striving behavior. However, beyond the clear indication that the motivational forces in play *are* dynamic, the functional form of the relationships among goals, motivation, and performance over time, especially in the presence of a deadline, is not well understood.

The d statistic is used as a measure of effect size across studies in terms of the number of standard deviations by which the mean impact is shifted. For an intervention with $d = 0.52$, a task with expected duration of 10 minutes and a standard deviation of 2 minutes would be expected to finish in $10 - 0.52(2) = 8.96$ minutes. This paper *assumes* a level of improvement from goal setting and builds a performance model and procedure from that assumption. Then, it investigates the resulting performance improvements, presuming the goal-setting intervention works. After the primary results are presented, sensitivity analysis is conducted on the size of the goal-setting effect, including what would happen if the intervention fails.

In addition to empirical work, there has been some theoretical modeling of work rate changes over time, which might be explained by a motivational dynamic. For example, Graves (1986) modeled a work rate that increased conjointly with workload, while Gutierrez and Kouvelis (1991) modeled a work rate that declined as the time to a deadline extended. These models may be useful, but they do not rest on an explanatory base; there is no empirical work that supports why work rates should change in the way these models suggest. One exception to this is the work of Schultz et al. (1998), which drew on substantial support from the Organizational Behavior literature on, for example, interdependence to predict work rate changes due to variance in buffers. The model this paper presents is somewhat simpler, but more consistent with the limited state of empirical work on the dynamics of the goal-setting effect on motivation.

Goal difficulty is a variable that has received even more attention than deadlines. One method of operationalization that is comparable across studies and thus useful for meta-analysis is to use the frequency of goal attainment as a measure of difficulty (Wright 1990, Wright et al. 1995). To our knowledge, there is no single prescribed number or equation for a "difficult" vs. an "easy" goal. For example, one article referred to goals achieved 15% of the time or less as "difficult" and a goal achieved 50% of the time or more as "easy" (Klein et al. 1999), but prescriptions for setting a "difficult" goal have been given as low as 10% attainment and as high as 25% attainment. Nonetheless, it is clear that if a goal is perceived as "easy," it loses its motivational force, and if perceived as unrealistically difficult, it is simply rejected and again loses its motivational force (Erez and Zidon 1984).

Of course, what is easy or hard will itself be affected by motivation, and this suggests the need for a dynamic model in which changes in motivation and goal difficulty interact. Such effects would be particularly important to capture in the intermediate or long term. Early work by Ivancevich (1972) suggested that the motivational effects of a goal change over the long term and might be maintained only if reinforced. However, the dynamics remain under-investigated, and our approach represents only a small step toward addressing this issue.

The presence of a deadline adds another dimension to the issue of goal difficulty. A task arriving closer to a deadline will be more difficult to finish than one arriving earlier. A precise level of attainment cannot be applied to any given task—some tasks must necessarily be harder than others, and tasks which start out being "easy" may end up being "difficult." One could simply set a new goal for each task as it arrives. However, given the information processing burden of this approach, setting a stationary goal based solely on information available before the task arrives is the preferable method. How to make such a goal motivational in aggregate for all the tasks is still an open question.

In the context of our field setting, managers are especially sensitive to time-based measures that relate to customer service. In reviewing and evaluating the metrics used by logistics managers, Caplice and Sheffi (1994) said that a metric assessed the effectiveness of a logistics operation when the metric compared output with some normative standard. Of the 13 common effectiveness metrics they list, 8 used a time period as the standard. The connection between the timeliness of operations and customer service is well understood in logistics and warehousing. Johnson and Davis (1998) describe how Hewlett-Packard tracks the timeliness of order fulfillment in logistics operations by examining not only the average on-time performance, but the entire distribution of order fulfillment times against their deadlines, or customer-promised dates (which they call *order aging profiles*).

The NSD metric used in this paper is a time-based effectiveness measure, closely related to order aging profiles. The performance model presented in the next section shows how a motivational goal on the NSD metric can change performance. As such, it begins to fill the gap between the findings of Goal-Setting Theory and the operational metrics and policies used to monitor and control order fulfillment operations.

3. Performance Model

In this section, a stochastic model of completion times is developed for a randomly arriving task, when a time-sensitive goal may motivate performance. This model is required to evaluate the effect of a motivational goal on a deadline-sensitive task, when the task may arrive early or finish late. In the next section, we will embed this model in a Monte Carlo simulation to evaluate potential cutoff times and nominal (percent-to-complete) targets.

The model of behavioral effects presented here is relatively simple because there is no empirical evidence to guide the development of a more sophisticated model. Nevertheless, our results suggest that a motivational goal, coupled with appropriate parameters for the NSD policy, can improve performance as seen by the customer.

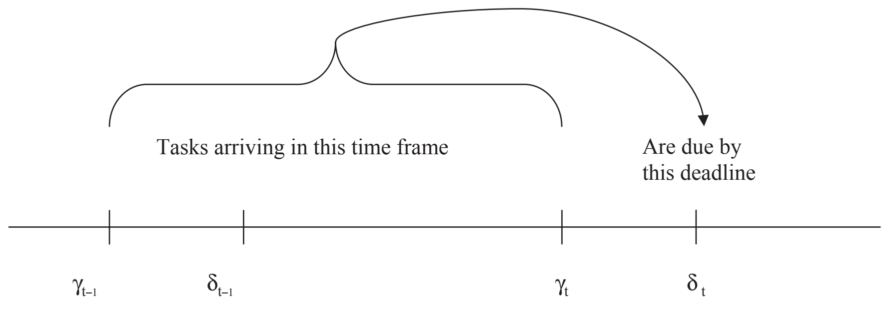
3.1. A Model of Completion Times with Motivated Workers

Consider an order fulfillment operation with orders arriving in a continuous stream throughout the day. Workers prepare orders and make them available for shipping (see Figure 1). Orders arrive in a work cycle t. A work cycle has both a cutoff time and a targeted shipment (truck departure) time. An order arriving before a cutoff time is targeted for shipment on the next truck scheduled to depart after that cutoff time.

The processing time of an order is defined as the time between its arrival and the time it is available for shipping. Note that this definition of processing time includes both operating time and waiting time in queues. The essence of the model is that motivated workers find ways to make orders available for shipping more quickly. In a later section, simulation is used to investigate the sensitivity of our model to this assumption, to investigate the impact of queuing and congestion.

Performance on the NSD metric depends on the distribution of arrival times and processing times. It is assumed that the firm has historical data on orders, including arrival times and processing times for "unmotivated workers," or workers not responding to NSD or some other motivational metric. The arrival and processing time distributions convolve to determine a distribution *(i.e., density function)* of completion times $f_c(x, y)$, where completion time of an order is the

Figure 1 Cutoff Times and Deadlines



sum of its arrival time x and processing time y. Orders arriving before a specified cutoff time γ_t each day are batched and delivered to customers on a departing vehicle. Assuming that there are no motivational effects, expected performance on the NSD metric is

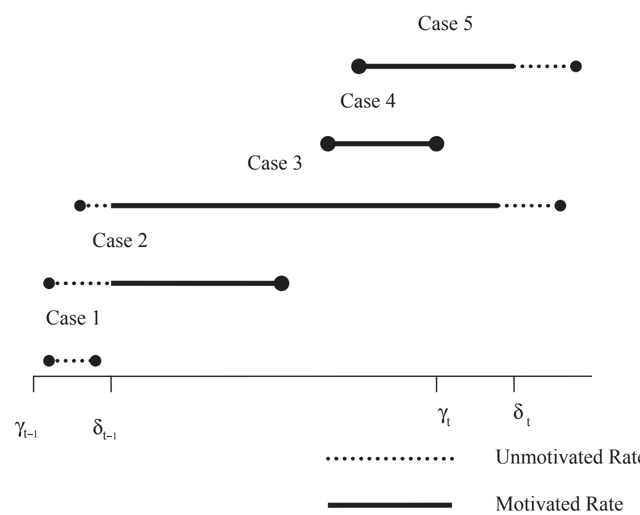
$$N_u = \int_{x=\gamma_{t-1}}^{\gamma} \int_{y=0}^{\delta_t - x} f_c(x, y)dydx, \qquad (1)$$

where δ_t is the target shipment time for the cycle and γ_t is the cutoff time for jobs/arrivals counted within a cycle. In other words, we are measuring the probability that a random order (arriving between two cutoff times) will *finish by the deadline*. Because order arrival and processing times are independent and identically distributed, the probability that a randomly arriving order finishes before the deadline can be used as a surrogate for the proportion of all orders that finish. With this expression for the unmotivated completion time distribution and expected NSD performance, we can apply some of the findings from Goal-Setting Theory to establish the cutoff time.

Recall that a goal is said to motivate an employee on a task in part through a mechanism that involves attention. The goal directs attention to certain tasks at the expense of others. Once attention is focused on goal-relevant tasks, one can expect persistence and effort to increase and reduce processing time. For example, a task arriving after the cutoff but before the previous deadline cannot be expected to receive the same attention as the tasks that need to be completed for the more proximal goal. Moreover, given random arrivals across the deadline periods and a relatively long processing time, there is a significant chance that some tasks will not be finished by the deadline. In that case, a new goal becomes relevant, and work on the incomplete tasks can no longer help attain that new goal.

This leads to the definition of a *motivated window* of time [δ_{t-1}, δ_t] for tasks that arrive in the interval [γ_{t-1}, γ_t]. A task that arrives between [γ_{t-1}, δ_{t-1}] will be worked at an unmotivated rate until δ_{t-1}, because the task is not relevant to the most proximal goal. A task that arrived before γ_t but is unfinished by δ_t will no longer be worked at the motivated rate, because it is no longer relevant for attaining the goal. Figure 2 shows the five possible combinations of motivated and unmotivated task work rates.

Figure 2 Motivated Window



The motivated rate itself can be expressed in terms of the expected motivational effect of goal setting. Recall from the previous section that this is expressed in terms of the meta-analysis coefficient d, that is, as a reduction in task time equivalent to d standard deviations. Given our expected unmotivated processing time,

$$\bar{t}_u = \int y f_p(y)dy,$$

the expected motivated processing time will be

$$\bar{t}_m = \bar{t}_u - d\sigma_p,$$

where σ_p is the standard deviation of $f_p(y)$. The corresponding rates are

$$r_u = \frac{1}{\bar{t}_u}, \qquad (2)$$

and

$$r_m = \frac{1}{\bar{t}_u - d\sigma_p}. \quad (3)$$

As r_m increases, the processing time distribution $f_p(y)$ will underestimate the probability that a task will finish within a given time period and hence (1) is an underestimate of NSD performance, given the presence of a motivated window.

Equations (2) and (3) can be used to develop formulae for task completion time z for each of the five cases shown in Figure 2. The expressions in Table 1 rely on the fact that the duration of motivated tasks (or motivated portions of tasks) is proportionally (r_u/r_m) less than the duration of unmotivated tasks, or conversely, that (r_m/r_u) times more units of work can be accomplished at the motivated rate. The expressions for Case 1 and Case 4 follow directly from this observation.

In Case 2, the motivated portion of the task starts at δ_{t-1}, so the portion of t_u that is worked at the unmotivated rate is $\delta_{t-1} - x$, where x is the arrival time. The remainder of the task time is

$$t_u - \delta_{t-1} - x,$$

which will be completed in $(r_u/r_m)[t_u - \delta_{t-1} - x]$ time units.

In Case 5, the task arrives in the motivated window but has not been finished by time δ_t so the remaining work must be finished at the unmotivated rate. The remaining work is

$$t_u - \{\delta_t - x\}(r_m/r_u),$$

because in the time $\{\delta_t - x\}$, work was accomplished at the motivated rate.

In Case 3, we have the combination of Cases 2 and 5: unmotivated work at the beginning and end of the task. As in Case 2, the first unmotivated portion ends at time δ_{t-1}, and has finished $(\delta_{t-1} - \mathrm{x})$ time units of work. So, at δ_{t-1},

$$t_u - \delta_{t-1} - x$$

time units of work remain. In this case, motivated work will proceed for the entire cycle, accomplishing

$$(r_m/r_u)[\delta_t - \delta_{t-1}]$$

units of work. Hence, the remaining work at time δ is

$$t_u - (\delta_{t-1} - x) - (r_m/r_u)[\delta_t - \delta_{t-1}].$$

Recall that the nominal target is stated by management as a goal for NSD: the percentage of orders that workers should attempt to place on the targeted truck. Given a cutoff time γ_t, these expressions can be used to predict motivated NSD performance N_m against a nominal target π by using a Monte Carlo simulation to integrate the cases across the joint distribution of arrival times and unmotivated processing times. But we have not yet dealt with the problem of how to determine values for γ or π.

4. A Goal-Setting Procedure

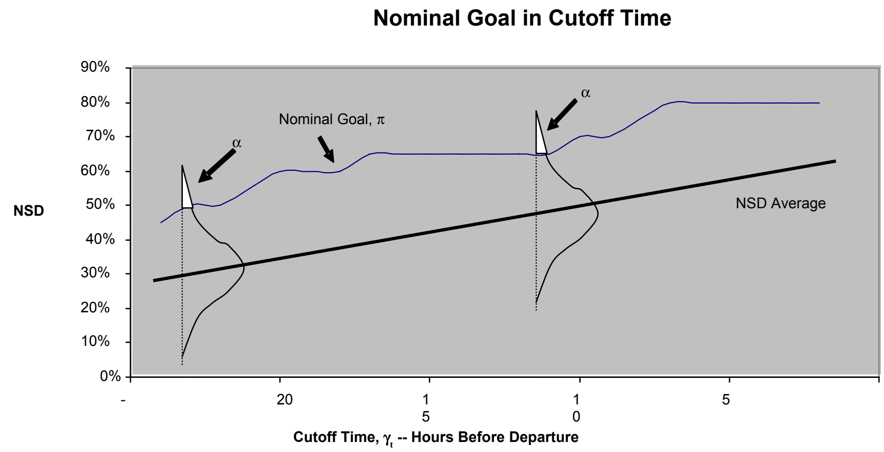
In the goal-setting literature, a motivational goal needs to be difficult, which is usually defined in terms of α, the frequency of goal attainment. A difficult goal is one that has a relatively small chance of being attained (e.g., α = 20%). However, having decided on a level of goal attainment, there is still more than one way to set π and γ to create a policy that has the desired level of difficulty. To see this, consider that a given nominal goal π will be more easily attained by moving the cutoff time γ_t backward in time. Conversely, a given γ_t can be made to correspond to a relatively easier goal by lowering π. The relationships among π, γ_t, and α are shown in Figure 3.

In Figure 3, a different distribution of N_m is shown for different values of γ_t. On each of those distributions, π is the quantile associated with a performance that is obtained only α% of the time. The distribution of N_m is a function of the (motivated) distribution of task completion times given in Table 1, and π is a point on the distribution of N_m that requires estimation. Note that the two example distributions drawn there are merely notional—the true distributions are not necessarily symmetric or identical (and the field-site data are neither). Whether or not the distributions are symmetric or identical, generally α will be a

Table 1 A Model of Motivated Task Completion Times Against a Deadline (Time Measured in Days, Setting Previous Cutoff Time = (γ − 1) = 0)

Case	Arrival time condition	Processing time condition	Completion time
1	$\gamma_{t-1} < x < \delta_{t-1}$	$t_u < \{\delta_{t\text{-}1} - x\}$	$z = x + t_u$
2	$\gamma_{t-1} < x < \delta_{t-1}$	$(t_u - \{\delta_{t-1} - x\})(\frac{r_u}{r_m}) < (\delta_t - \delta_{t-1})$	$z = \delta_{t-1} + (t_u - \{\delta_{t-1} - x\})\ (\frac{r_u}{r_m})$
3	$\gamma_{t-1} < x < \delta_{t-1}$	$(t_u - \{\delta_{t-1} - x\})(\frac{r_u}{r_m}) > (\delta_t - \delta_{t-1})$	$z = \delta_t + t_u - \{\delta_{t-1} - x\} - \frac{r_m}{r_u}\ (\delta_t - \delta_{t-1})$
4	$\delta_{t-1} < x < \gamma_t$	$t_u\ (\frac{r_u}{r_m}) < \delta_t - x$	$z = x + t_u\ (\frac{r_u}{r_m})$
5	$\delta_{t-1} < x < \gamma_t$	$t_u\ (\frac{r_u}{r_m}) > \delta_t - x$	$z = \delta_t + (t_u - [\{\delta_t - x\}\ (\frac{r_m}{r_u})])$

Figure 3 Relationships among γ, π, and α



different point on *each* distribution that corresponds to a nominal goal that achieves the desired level of goal difficulty. Because the Central Limit Theorem does not apply to quantile estimates on these distributions, a procedure such as bootstrapping (Efron and Tibshirani 1998) must be used to provide robust estimates of π. (Note that this bootstrapping procedure would not be needed if the task-completion time distributions could be fit to a simple analytical distribution. However, because we are estimating the distribution from empirical arrival and processing time distributions, it is necessary.)

Goal-setting theory prescribes not only a level for α, but a level for γ_t as well. In Figure 2, for example, the motivational window is maximized when $\delta_t - \gamma_t$ is as small as possible. To make $\delta_t - \gamma_t$ as small as possible, one should make γ_t as late as possible, then one can use Monte Carlo simulation with Equation (1) and the equations in Table 1 to set π to the nominal goal, which can be achieved α% of the time. Because it provides a motivational goal over the largest motivational window, this set of policy parameters should yield the highest customer service.

In practice, an upper bound on γ_t might be determined by three factors. First, one must worry about goal rejection. For example, if average processing time is 12 hours, a cutoff time set to 1 hour before departure is likely to be rejected as unrealistic. We are unaware of any goal-setting research to guide the setting of a bound to guarantee goal acceptance. The implications of this will be examined below. Second, the magnitude of the nominal goal may become so low that it lacks face validity as a stretch goal for the workers. For example, telling a worker that he must "strive to ship 30%" of the available orders on the next scheduled departure is unlikely to be perceived as a difficult goal, and workers' failure to achieve such a seemingly modest goal more than 20% of the time is likely instead to be de-motivational. Finally, if a cutoff time is published to the customers, customer service concerns will dictate an earlier cutoff time. In our procedure, we do not bound γ, but allow it to be as late as one increment (1 hour) before the truck departure. An arbitrary constraint might be added to our procedure, to ensure a minimum time period between the cutoff time and truck departure. We will examine the practical implications of so bounding the cutoff time in the next section.

5. Application and Analysis

In this section, data from the field site previously mentioned are used to demonstrate the procedure used to set goal parameters. Using the tenets of Goal-Setting Theory, α is set to a "difficult goal" level, and γ_t is made as late as possible to maximize the size of the motivated window. Given these values for α and γ_t, a Monte Carlo simulation and bootstrapping procedure are used to estimate π and to calculate the expected performance of the policy. Next, our results are compared to a policy that management found intuitively appealing, and the superiority of our approach on these data is demonstrated. Then, a sensitivity analysis is conducted to assess the effect of setting an upper bound on cutoff time or a lower bound on the magnitude of the nominal goal. Finally, discrete-event simulation is used to investigate the impact of congestion effects on our results and to investigate the sensitivity of our results to changes in effect size.

5.1. Applying the Policy

To apply the policy, we must assume a particular value for the goal-setting effect. For our numerical analysis, two moderators of this effect are particularly relevant. One is the difference between field applications and lab studies, with meta-analyses of field applications typically reporting smaller effect sizes. For example, Mento et al. (1987) reported an effect size of $d = 0.439$ in field studies, but $d = 0.624$ in lab studies. Likewise, Tubbs (1986) reported an effect size of $d = 0.520$ in the field, but 0.897 in the lab. Another moderator that is important to us is task complexity, with simpler tasks showing a stronger effect in meta-analyses. For example, Wood et al. (1987) report performance improvements of 12.15% ($d = 0.76$) on simpler tasks, but only 7.8% on more complex tasks. In the present study, an effect size of $d = 0.52$ is assumed as a realistic but conservative estimate, recognizing that the exact effect will vary from setting to setting. The paper will later examine other effect sizes (including no effect) in a *post hoc* analysis.

The sample data consisted of the 9,133 observations of arrival and processing times from 1 month for a single transportation provider, an average of 295 orders/day. The site had several providers, as well as its own fleet of transportation assets. The shipments we chose to track had a reliable daily shipment schedule, provided by a well-known third-party logistics provider. Mean processing time was 16.48 hours, with a standard deviation of 10.07 hours. Both the processing time and arrival time distributions were skewed and multi-modal. No analytical distribution could be found to provide an adequate fit to either, so we chose to use the empirical distributions in our simulation.

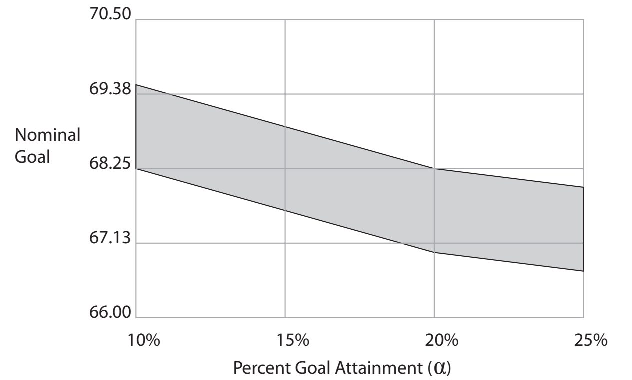
To maximize the size of the motivated window, the cutoff time was set to 1 hour before the deadline. A Monte Carlo simulation and the equations in Table 1 were then used to derive a finish time distribution for each task, accounting for motivation effects. We simulated 300 tasks for each deadline, roughly the average number of tasks per day at our site, and calculated the percentage of tasks that finished by the deadline, N_m. Each run of this simulation provides us with a single point on a sample distribution of N_m for that cutoff time. The simulation was then repeated 100 times, yielding a sampling distribution of 100 observations of performance for each cutoff time. The CDF of this distribution is $\hat{\Phi}_\gamma(N_m)$, and $\pi = N_m \in \hat{\Phi}_\gamma(N_m) = \alpha$, that is,

$$\hat{\Phi}_\gamma^{-1}(\alpha) = \pi. \tag{4}$$

We set $\alpha = 0.2$, and so find $\hat{\Phi}_1^{-1}(0.2) = \pi$.

Note that $\hat{\Phi}_\gamma(N_m)$ is the sampling distribution of a point estimate, the percentile associated with the number of tasks completed by the deadline. To assess the quality (confidence) of our estimate, a bootstrap analysis was conducted on our Monte Carlo simulation results, resampling $\hat{\Phi}_\gamma(N_m)$ 500 times to get acceptably low variance in the estimate (Efron and Tibshirani 1998, p. 275). A 90% confidence interval is shown in Figure 4, which indicates that the nominal goal should be set to 68%. Importantly, Figure 4 also shows that, on these data at least, the procedure is relatively insensitive (robust) to the particular value selected for α. This is desirable because there is no precise guidance on the best value for α.

Figure 4 Ninety Percent Confidence Intervals for the Nominal Goal, with Cutoff Time 1 Hour before Departure, across Values of Goal Attainment from 0.10 to 0.25



5.2. Performance Comparisons

To assess the quality of the policy derived on these data ($\alpha = 0.2$, $\gamma_t = 1$ hour before departure, $\pi = 0.68$), it is next compared with two heuristic policies proposed by managers at the industry site. To make this comparison, there must be a common yardstick, separate from NSD, against which to compare the policies. As one varies cutoff time and nominal goal, NSD has different meanings—comparing a policy of $\gamma_t = 23$ hours before departure, $\pi = 50\%$ (that is, up to 23 hours to get 50% of the tasks finished) with a policy of $\gamma_t = 1$ hour before departure, $\pi = 90\%$ (that is, as little as 1 hour to get 90% of the tasks finished) in terms of percent of tasks finished by the deadline is not a fair comparison. This problem was solved by measuring the percentage of tasks that were finished by the deadline immediately following arrival, *regardless* of whether that deadline was *targeted* for the task. This measure is related to the order aging profile metric mentioned in Johnson and Davis (1998) and can be called "tasks finished without delay." The percentage of tasks finished without delay under our policy was 61.3%.

Management at our field site intuitively felt that the nominal goal π should be some high number, such as 80%, to be motivational. They proposed a policy (A) to use $f_p(y)$ to determine the processing time required by at least 80% of the tasks and to set the cutoff time back from the deadline by that many hours—in this case, 21 hours before the deadline. Setting values for π and γ_t, of course, determines a value for α. One might think that this would yield an "easy" goal that would be attained about 50% of the time, but that is not necessarily the case, because setting the parameters in this way ignores the effect of the arrival time distribution. However, when we simulated this policy with 300 tasks over 100 deadlines, we found that the nominal goal was indeed attained 100% of the time (even without a motivational effect to reduce processing times). The minimum percentage of tasks completed by the deadline was 88.67% and the maximum was 95.67%. However, percentage of tasks finished without delay was only 41.1%, primarily due to the fact that this easy goal was not motivational.

The failure of policy A suggests policy B: set a nominal goal of 80%, but then use Equation (4) to search over cutoff times until one finds a cutoff time that yields a difficult nominal goal of 80%. While this procedure is less intuitive, it still relies on the model developed in this paper. Policy B can be stated as

$$\gamma_t \hat{\Phi}_\gamma^{-1}(0.2) = 0.8. \tag{5}$$

Figure 5 illustrates the result of this search. In Figure 5, the dashed line shows the nominal goal at each cutoff time, while the solid line displays the percentage of orders finished without delay—our common yardstick performance measure. The cutoff time most nearly associated with a difficult goal of 80% is 10 hours before departure. The percent of tasks finished without delay under this policy is only 58.2%, slightly worse than the 61.3% obtained by our policy. As Figure 5 demonstrates, the best policy is one that maximizes the size of the motivational window by setting the cutoff time as late as is practical.

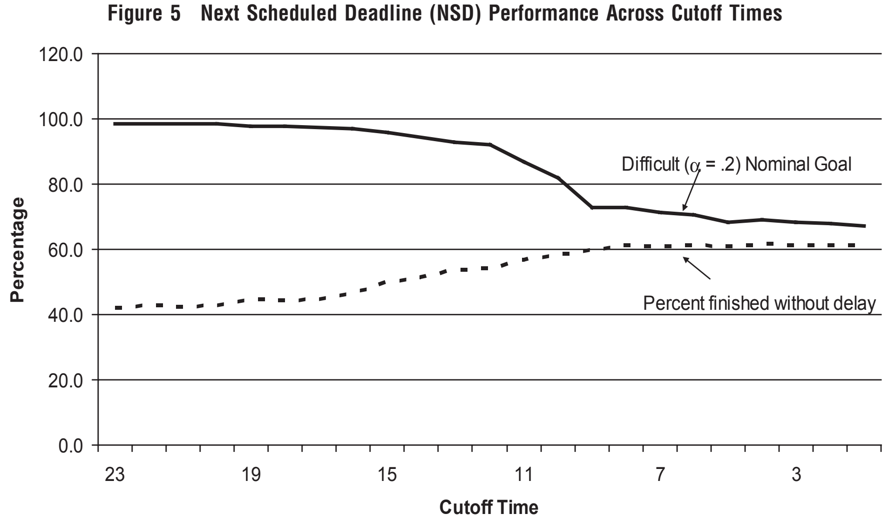
In spite of the implications of Figure 5, there may still be practical reasons to use Policy B, or a policy that bounds cutoff time rather than the nominal goal. These will be discussed in the next section. The results show that the price of establishing a "floor" such as 80% on the nominal goal percentage or a ceiling on cutoff time such as 10 hours before departure time may be quite low, so long as the proper goal-setting procedure is used.

5.3. Sensitivity Analysis and Effect of Congestion

Before addressing the managerial implications of the model, consider two ancillary issues—the sensitivity of our results to changes in effect size (d) and the effect of congestion. The goal-setting effect size used in the main analysis was $d = 0.52$. Again, the results *assume* this goal-setting improvement and build a performance model from it. Based on that performance model, a procedure for setting the cutoff time and nominal goal was proposed consistent with the tenets of Goal-Setting Theory. However, goal-setting interventions do *not* always work. And while $d = 0.52$ may be a fair representation of the average effect size to be expected in a field implementation with simple tasks, the effect size in any particular setting can be more or less. To address this issue, a discrete-event simulation was developed to examine effect sizes across a range from 0.0 (failed implementation) to $d = 0.75$ (a conservative "maximum" that might occur in a field implementation). We assumed a simple, single-server queuing system with exponential processing and interarrival times, adjusted to achieve a utilization of approximately 85 percent. The cutoff time was held constant at 4 hours before truck departure.

For effect size $d = 0$ (no motivation), average NSD was 84.2 percent. The simulation was repeated for effect sizes $d = 0.1$, 0.25, 0.52, and 0.75. Average NSD was 92.8, 95.8, 99.2, and 99.9 percent, respectively. As expected, performance is improved, even for quite modest effect sizes, so long as the goal-setting intervention does not actually fail. Incidentally, one might still expect an improvement from the NSD metric

Figure 5 Next Scheduled Deadline (NSD) Performance Across Cutoff Times



compared to the flow time metric even without the motivational effect, because workers will no longer be working overtime on Friday to put boxes on the dock to be shipped Monday afternoon. However, these effects are not captured in the simulation.

The performance model ignores the effects of congestion, which may seem curious to those familiar with outbound operations in a warehouse. Because our goal is set in terms of when work is started rather than when an order arrives, the primary effect of congestion will not be on motivation. At the field site, orders were not released to workers until they were ready to start them. If workers can see the work "build up," congestion may have motivational effects as well (Schultz et al. 1998, 1999). However, in periods of high utilization, there will be a delay in *starting* work, and hence, the percentage of tasks *finished* without delay will be reduced as utilization increases. To demonstrate this, a discrete event simulation was coded and average utilization was varied from 55% to 98%. Results are shown in Table 2. The percent reduction in task times is related to the goal effect size d. As is shown, congestion has relatively little effect on motivation. In heavy traffic, however, our model's predictions for tasks finished without delay are significantly overstated. The prescription remains the same, regardless of utilization: set cutoff times as late as practical, but heavy traffic will reduce the performance impact of the motivational goal.

6. Discussion and Conclusions

Goal acceptance and commitment have previously been investigated as moderators (Klein et al. 1999, Ronan et al. 1973), but beyond the general result that goals can be too difficult, there is little specific enough to build a model upon. The assumed bound on cutoff time (at least 1 hour before the departure) is recognition of the importance of goal acceptance, which requires little in the way of additional assumptions. More sophisticated approaches are possible, even given the limited data available about precisely when goals will be rejected. For example, another approach to bounding the cutoff time would be to examine the processing time distribution to find the minimum time required by at least 20% of the orders. In the field data, this was 8.1 hours. Setting the cutoff time at least 8 hours before the goal ought to give employees at least a 20% chance to finish *any* task that arrived before the cutoff time. As 20% has been shown to yield a difficult goal that is not rejected, this should provide an acceptable upper bound. As Figure 5 indicates, a cutoff time 8 hours before the deadline yields only a very small degradation in performance on our field data (tasks finished without delay drops to 59.8 from 61.3). Ultimately, however, there is no guarantee that any goal will be accepted. Good scales exist, and goal acceptance is best addressed by measuring it.

Table 2 Effects of Congestion

Utilization %	% Reduction in task time	Average queue delay (hours)	% of tasks finished without delay
55	24.3	0.0002	53.3
65	24.3	0.0033	53.4
75	24.2	0.0370	53.2
85	24.0	0.3463	51.9
95	22.0	6.7996	30.0
98	21.2	25.1800	17.9

Note that the percentage tasks finished without delay do not match the numbers reported in Figure 5 because analytical distributions were used, rather than the empirical distributions used in Figure 5. We cross-validated the results shown here by applying the analytical distributions back into the model shown in Table 1.

A more difficult question is whether or not to publish cutoff times to customers. This is done in some distribution businesses, for example, when customers are told that they will receive their order the next day if they place it the previous day by 14:00. In such cases, what is purely a nominal goal in our model becomes the percentage of time a customer-promise date is fulfilled, and consequently values of π as low as 68% are unacceptable. On the other hand, cutoff times set too early will not provide any marketing advantage (customers are likely to be unimpressed if told that their order will arrive the next day, so long as they order by 1:00 a.m. on the previous day). The development of a model to capture these customer service costs is beyond the scope of the current paper, but Figure 5 makes some of the trade-offs plain. On the data from the field site, there is very little reason to move the cutoff time earlier than 16 hours before the deadline, where the nominal goal is 97% and the percentage of tasks finished with delay is 46.5%. Cutoff times set any earlier yield minimal improvements in the nominal goal at a large expense in the percentage of tasks finished without delay (down to 41.9%), as well as the size of the motivational window and the "marketing opportunity" of a late cutoff time. On the other hand, after $\gamma_t = 12$ hours before the deadline, the nominal goal drops off quite rapidly, moving from 92% down to 73% only 3 hours later. Setting γ_t to between 16 and 12 hours before the cutoff, the choice of a cutoff time would depend on the relative marketing value of the later cutoff time, against the cost of lowered delivery-as-promised from 97% to 92%. These times are wholly a function of the particular arrival and processing time distributions at our field site. However, given other arrival and processing time distributions, the procedure will still yield appropriate (motivational) nominal goals for each cutoff time, so that management can make a better informed decision about setting and publishing the cutoff time.

A simpler method might simply be to establish two cutoff times—there is no reason why the cutoff time given to motivate employees must be the same as the cutoff time published to customers. The intent of each is different. The cutoff time would have to be set early enough to get nearly all the orders that arrive before the customer-published cutoff time on the next shipment (i.e., the customer-published cutoff time should yield an "easy" and hence non-motivational goal). The manager of an order fulfillment operation may object that there is nothing easy about meeting customer expectations for nearly perfect on-time delivery. Partly this is just confusion raised because of the way the word "easy" has been defined in the goal-setting literature—it is not "easy" to meet any performance target consistently and reliably.

Finally, the NSD metric allows executives of large distribution firms with multiple, disparate distribution centers to compare service performance fairly. Because NSD can be "tuned" with the cutoff time, executives can establish base cutoff times that correspond to the same nominal goal for all DCs, thereby establishing a common yardstick for performance. DC managers at our field site occasionally complained that measures used to judge their performance against other DCs were, necessarily, "apples and oranges." NSD provides a way to move toward "apples and apples."

6.1. Limitations

Before summarizing the contributions of this research, we review its limitations. The first concerns the external validity or generalizability of the results. The findings reported here are based on a simulation model of the process in question and not on experimentation with the process itself. Although the use of simulation models to investigate behavioral phenomenon is becoming increasingly common and more widely accepted (e.g., Vancouver et al. 2005), such research will always need further cross-validation from field and laboratory work. The field data, the particular values of the parameters derived (i.e., π and γ_t), as well as the performance improvement reported all come from a unique setting. A particular concern is that the coefficient of variation (CV) in the processing times examined is unusually high (0.61) compared to typical ranges of CV found with unpaced tasks (Doerr and Arreola-Risa 2000, Knott and Sury 1987, Muth 1973). For example, Doerr et al. (2000) characterize "high variability" unpaced tasks as those in the range of 0.22–0.57. Because the magnitude of the goal-setting effect depends on the relative magnitude of the standard deviation to the mean (i.e., the CV), the magnitude of the performance improvement reported here may be greater than that which would be found in other field settings. However, having demonstrated that the performance model and bootstrapping procedure is robust enough to yield good results with these unusual distributions, there is no reason to suppose that they could not be even more easily applied to a situation where tasks and arrivals followed a more predictable pattern.

Another limitation is that the procedure assumes that the goal intervention will succeed, but that is not always the case. Still, goal-setting interventions are widely prescribed and widely successful: the effect sizes reported here are *average* effect sizes from meta-analytic studies. Sensitivity analysis on effect size has shown that significant performance improvement is obtainable with quite modest effect sizes. Figure 4 provides a limited sensitivity analysis on α, showing that the procedure is robust across reasonable interpretations of a "difficult goal." This provides partial evidence that the bootstrapping procedure (Equation 4) will provide a motivational goal, regardless of the exact value of α, so long as α is chosen to be a level that is neither too easy nor too hard. We recognize the limitations of these assumptions, but if the empirical results of goal-setting research are ever to be applied in modeling work, similar assumptions will have to be made, and we are not aware of any more rigorous procedure to select a value for d or α.

Third, the basic performance model ignores congestion effects. As the sensitivity analysis shows, the primary impact of queuing delay is not on motivation. However, when utilization is high, significant order delays will occur, and the basic model will be optimistic in terms of the expected number of tasks finished without delay.

Finally, although the performance model is a fair representation of the current state of goal-setting effects in a dynamic environment, better models are needed for both near- and intermediate-term policy making. In the intermediate term, the feedback loop implied by the goal setting–motivation interaction needs to be modeled explicitly. The "motivational windows" defined here are only a first-stage approximation to this interaction, for which a system-dynamic model might be developed. In the near term as well, our policy does not use the sorts of information technology available today to modify goals based on information about orders and operations. There is, however, currently no clear guidance on what to do with such information in terms of setting goals. Important extensions to Goal-Setting Theory are needed to investigate exactly how motivation shifts over time in the near term, especially as a deadline approaches. The model in this paper, in which motivation is either present or absent depending on the current time in relation to the targeted deadline, could be enhanced were such information available.

6.2. Summary and Conclusions

For order fulfillment systems with deadlines, a sensible performance metric for customer service must incorporate the linkage between upstream continuous processes and downstream batch processes. The NSD metric does just this, and in a way that allows managers to adjust to any order arrival or processing time distribution. When published to workers along with a properly established goal, the metric encourages motivated behaviors such as increased work rates and improved task strategies. This motivation in turn improves customer service by causing more customers to receive their orders sooner. An important feature of the metric is that, for a given cutoff time, an increase in its recorded value necessarily means an improvement in real customer service, which in this case is reduced customer waiting time.

The procedure takes as input the distributions of task arrival and processing times. Using principles from Goal-Setting Theory, a performance model (Table 1) is developed to set a cutoff time between deadlines. Tasks arriving after the cutoff time are not targeted for the immediately following deadline, but for the subsequent one. To set a motivational goal for each deadline, principles from Goal-Setting Theory are again applied to determine a desired level of difficulty, then a bootstrapping procedure is applied to estimate a nominal goal corresponding to that level of difficulty for the given cutoff time.

The policy resulting from the procedure is compared with an intuitive policy, using the percentage of tasks finished without delay as a criterion. The procedure was demonstrated to be superior on these data. Under the procedure, 61.3% of tasks were finished without delay, compared to 41.1% under the intuitive policy. Also investigated was the penalty that would be paid if management established a floor on the nominal goal or a ceiling on the cutoff time. Monte Carlo simulation results suggested that this penalty might be as low as 3.1% (58.2% of tasks finished without delay, as opposed to 61.3%).

References

Bendoly, E., R. Croson, P. Goncalves, K. Schultz. 2010. Bodies of knowledge in behavioral operations. *Prod. Oper. Manag.* **19**(4): 434–452.

Bendoly, E., M. Prietula. 2008. In the zone: The role of evolving skill and transitional workload on motivation and realized performance in operational tasks. *Int. J. Oper. Prod. Manag.* **28**(12): 1130–1152.

Boudreau, J., W. J. Hopp, J. O. McClain, L. J. Thomas. 2003. On the interface between operations and human resources management. *Manuf. Serv. Oper. Manag.* **5**(3): 179–202.

Caplice, C., Y. Sheffi. 1994. A review and evaluation of logistics metrics. *Int. J. Logist. Manag.* **5**(2): 11–28.

Doerr, K. H., A. Arreola-Risa. 2000. A worker-based approach for modeling variability in task completion times. *IIE Trans.* **32**(7): 625–636.

Doerr, K. H., T. D. Klastorin, M. J. Magazine. 2000. Synchronous unpaced flow lines with worker differences and overtime cost. *Manage. Sci.* **46**(3): 421–435.

Doerr, K. H., T. R. Mitchell, T. D. Klastorin, K. A. Brown. 1996. Impact of material flow policies and goals on job outcomes. *J. Appl. Psychol.* **81**(2): 142–153.

Efron, B., R. J. Tibshirani. 1998. *An Introduction to the Bootstrap*. Chapman & Hall CRC, Boca Raton, FL.

Erez, M., I. Zidon. 1984. Effect of goal acceptance on the relationship of goal difficulty to performance. *J. Appl. Psychol.* **69**(1): 69–78.

Euske, K. J., L. A. Zander. 2005. History of performance measurement. K. Kempf-Leonard, ed. *The Encyclopedia of Social Measurement*. Elsevier, San Francisco, CA, 227–231.

Graves, S. 1986. A tactical planning model for a job shop. *Oper. Res.* **34**(4): 522–531.

Gutierrez, G. J., P. Kouvelis. 1991. Parkinson's law and its implications for project management. *Manage. Sci.* **37**(8): 990–1001.

Hays, J. M., A. V. Hill. 2001. A longitudinal empirical study of the effect of a service guarantee on employee motivation/vision, service learning, and perceived service quality. *Prod. Oper. Manag.* **10**(4): 405–424.

Hays, J. M., A. V. Hill. 2006. An extended longitudinal study of the effects of a service guarantee. *Prod. Oper. Manag.* **15**(1): 117–131.

Ivancevich, J. 1972. A longitudinal assessment of management by objectives. *Adm. Sci. Q.* **17**: 126–138.

Johnson, D. W., G. Maruyama, R. T. Johnson, D. Nelson, L. Skon. 1981. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychol. Bull.* **89**(1): 47–62.

Johnson, H. T., R. S. Kaplan. 1987. *Relevance Lost: The Rise and Fall of Management Accounting*. Harvard Business School Press, Boston, MA.

Johnson, M. E., T. Davis. 1998. Improving supply chain performance by using order fulfillment metrics. *Natl. Prod. Rev.* (Summer, 1998) **17**: 3–16.

Kerr, S. 1975. On the folly of rewarding A, while hoping for B. *Acad. Manag. J.* **18**(4): 769–783.

Klein, H. J., M. J. Wesson, J. R. Hollenbeck, B. J. Alge. 1999. Goal commitment and the goal-setting process: Conceptual clarification and empirical synthesis. *J. Appl. Psychol.* **84**(6): 885–896.

Knott, K., R. Sury. 1987. An exploratory study on stopping a paced line when incompletions occur. *IIE Trans.* **19**: 463–467.

Linderman, K., R. G. Schroeder, A. S. Choo. 2006. The role of goals in improvement teams. *J. Oper. Manag.* **24**(6): 779–787.

Linderman, K., R. G. Schroeder, S. Zaheer, A. S. Choo. 2003. Six-sigma, a goal-theoretic perspective. *J. Oper. Manag.* **21**(2): 193–203.

Locke, E. A., G. P. Latham. 1990. *A Theory of Goal Setting and Task Performance*. Prentice Hall, Englewood Cliffs, NJ.

Locke, E. A., G. P. Latham. 2009. Has goal setting gone wild, or have its attackers abandoned good scholarship? *Acad. Manag. Perspect.* **23**(1): 17–23.

Meier, G., M. H. Albrecht. 2003. The persistence process: Development of a stage model for goal-directed behavior. *J. Leadersh. Organ. Stud.* **10**(2): 43–54.

Mento, A. J., R. P. Steel, R. J. Karren. 1987. A meta-analytic study of the effects of goal setting on task performance. *Organ. Behav. Hum. Decis. Process.* **39**(1): 52–81.

Meyer, C. 1998. *How the Right Measures Help Teams Excel: Measuring Corporate Performance*. Harvard Business School Press, Boston.

Mitchell, T. R., D. Daniels. 2001. Motivation. W. C. Borman, D. R. Ilgen, R. J. Klimoski, eds. *Industrial and Organizational Psychology*. John Wiley & Sons, Inc., New York, 12.

Mitchell, T. R., W. S. Harman, T. W. Lee, D. Y. Lee. 2008. Self-regulation and multiple deadline goals. R. Kanfer, G. Chen, R. D. Pritchard, eds. *Work Motivation: Past, Present and Future*. Routledge Academic, Oxford, UK, 197–231.

Mitchell, T. R., H. Hopper, D. Daniels, J. George-Falvy, L. R. James. 1994. Predicting self-efficacy and performance during skill acquisition. *J. Appl. Psychol.* **79**(4): 506–517.

Muth, E. 1973. The production rate of a series of work stations with variable service times. *Int. J. Prod. Res.* **11**(2): 155–169.

Neely, A., R. Austin. 2002. Measuring performance: The operations perspective. A. Neely, ed. *Business Performance Measurement*. Cambridge University Press, Cambridge, 41–50.

Ronan, W. W., G. P. Latham, et al. 1973. Effects of goal setting and supervision on worker behavior in an industrial setting. *J. Appl. Psychol.* **58**(3): 302–307.

Schultz, K. L., D. C. Juran, J. W. Boudreau. 1999. The effects of low inventory on the development of productivity norms. *Manage. Sci.* **45**(12): 1664–1678.

Schultz, K. L., D. C. Juran, J. W. Boudreau, J. O. McClain, L. J. Thomas. 1998. Modeling and worker motivation in JIT production systems. *Manage. Sci.* **44**(12): 1595–1607.

Stalk, G., T. Hout. 1990. *Competing Against Time*. Free Press, New York.

Tubbs, M. E. 1986. Goal-setting: A meta-analytic examination of the empirical evidence. *J. Appl. Psychol.* **71**(3): 474–483.

Vancouver, J. B., D. J. Putka, C. A. Scherbaum. 2005. Testing a computational model of the goal-level effect: An example of a neglected methodology. *Organ. Res. Methods* **8**(1): 100–127.

Vancouver, J. B., C. M. Thompson, A. A. Williams. 2001. The changing signs in the relationships between self-efficacy, personal goals and performance. *J. Appl. Psychol.* **86**(4): 605–620.

Wood, R. E., A. J. Mento, E. A. Locke. 1987. Task complexity as a moderator of goal effects: A meta-analysis. *J. Appl. Psychol.* **72**(3): 416–440.

Wright, P. M. 1990. Operationalization of goal difficulty as a moderator of the goal difficulty-performance relationship. *J. Appl. Psychol.* **75**(3): 227–234.

Wright, P. M., J. R. Hollenbeck, S. Wolf, G. C. McMahan. 1995. The effects of varying goal difficulty operationalizations on goal setting outcomes and processes. *Organ. Behav. Hum. Decis. Process.* **61**(1): 28–43.

Yuen, G. 2006. Operation systems with discretionary task completions. *Manuf. Serv. Oper. Manag.* **8**(1): 114–117.